

# Polynomial Approximation of Functions in Sobolev Spaces\*

By Todd Dupont and Ridgway Scott\*\*

**Abstract.** Constructive proofs and several generalizations of approximation results of J. H. Bramble and S. R. Hilbert are presented. Using an averaged Taylor series, we represent a function as a polynomial plus a remainder. The remainder can be manipulated in many ways to give different types of bounds. Approximation of functions in fractional order Sobolev spaces is treated as well as the usual integer order spaces and several nonstandard Sobolev-like spaces.

**1. Introduction.** Approximation properties of finite element spaces are often derived using variations of the so-called Bramble-Hilbert Lemma [4], [5]. This lemma is based on an inequality of the form

$$(1.1) \quad \inf_{P \in \mathcal{P}} \|f - P\| \leq C \sum_{\alpha \in A} \left| \left( \frac{\partial}{\partial x} \right)^\alpha f \right|,$$

where  $\mathcal{P}$  is a class of polynomials,  $A$  is an associated class of multi-indices, and  $\|\cdot\|$  and  $|\cdot|$  denote certain Sobolev norms. An inequality of the form (1.1) can be found in Morrey [16] (and implicitly in Sobolev [18]) for the case of  $\mathcal{P}$  being all polynomials of degree at most  $r$ ,  $A$  being all multi-indices of length  $r + 1$ ,  $\|\cdot\|$  being the norm on  $W_p^m$ , and  $|\cdot|$  being the norm on  $L_p$ . In the second Bramble-Hilbert paper [5], (1.1) is derived for certain classes  $\mathcal{P}$  that range from the polynomials of degree at most  $r$  to the polynomials that are of degree at most  $r$  in each variable separately. Motivated by particular applications, we extend (1.1) (in Section 4) by allowing more general collections  $\mathcal{P}$  and  $A$  and, further, by deriving (in Section 5) inequalities of the form

$$(1.2) \quad \inf_{P \in \mathcal{P}} \|f - P\| \leq C \sum_j \left| P_j \left( \frac{\partial}{\partial x} \right) f \right|,$$

where  $\{P_j\}$  is a collection of homogeneous polynomials of degree  $l_j$  and  $\mathcal{P}$  is the intersection of the kernels of the operators  $P_j(\partial/\partial x)$ .

The proofs of Bramble and Hilbert used the results of Morrey and generalizations thereof. The proofs of these results are nonconstructive and cannot be used to estimate the size of  $C$  in (1.1) or to determine how  $C$  would vary as a function of the do-

---

Received August 23, 1978; revised December 12, 1978 and August 16, 1979.

AMS (MOS) subject classifications (1970). Primary 65D05; Secondary 41A63, 65N30.

\*Part of this research was done at Brookhaven National Laboratory under Contract No. EY-76-C-02-0016, U. S. Department of Energy. Part of the research of the second author was done at ICASE, NASA Langley Research Center under NASA Contract No. NAS1-14101.

\*\*Present address: Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109.

main. Sobolev's approach to imbedding theorems is based on an explicit representation of a function as a polynomial plus a remainder term. The results presented here use a related representation that is derived as an averaged Taylor's series (see Section 3). This representation can be manipulated in various ways to get bounds of the form (1.1) and (1.2). Although we do not explicitly calculate the associated constants here, it is easy to see what parameters they depend on, and, in particular cases, the proofs could be used to bound them. (The results of Section 5 are somewhat of an exception to this; see Remark 5.2.) We have calculated these constants in one special case [9]. Further, the form of proof used here allows the dependence of the constant on the underlying domain to be clarified. The basic results of this paper (Sections 3, 4, 5 and 6) are derived initially for domains that are star-shaped with respect to (each point of) a ball and in these cases the constants are seen to depend on the domain only through its diameter and the diameter of the associated ball. Having this type of dependence makes it easier (or possible) to treat the perturbations of the domain that are frequently needed to handle curved boundaries by finite element methods (see Example 4 in Section 8). Our results are also extended (in Section 7) to regions that may be viewed as a finite union of domains that are star-shaped with respect to balls. Polynomial approximation results for such regions have been derived by Jamet [14] using an entirely different approach. These regions include ones satisfying the cone condition used by Bramble and Hilbert [4], [5], but are slightly more general.

When functions are approximated by piecewise polynomials on a mesh of size  $h > 0$ , the bound for the error typically involves  $h$  to a positive power. In most cases, the power decreases by one for each additional order of differentiation applied to the error. One purpose of our results on tensor-product polynomial approximation (Section 4) is to show under what conditions one should expect *not* to lose a power of  $h$  when differentiating the error. An application is given to illustrate this point in Example 1 of Section 8.

There are situations in which it is necessary to approximate a function satisfying a homogeneous, constant coefficient differential equation by polynomials which also satisfy that equation. The approximation results following from (1.2) (see Section 5) can be used to treat such cases. An application is given in Example 2 of Section 8 in which harmonic functions are approximated by harmonic polynomials.

Our proofs of (1.1) and (1.2) are based on a basic representation formula of a function as a polynomial projection plus a remainder derived in Section 3. An important property of the projection operator is that it commutes with differentiation, that is, a derivative of the polynomial projection of a function is the same as an associated (lower order) polynomial projection of that derivative of the function. This commutativity property is used in a crucial way to derive the results described in the previous two paragraphs.

Frequently, one is interested in the best possible approximation of a function subject to the constraint that a function and its approximation agree at certain points [3], [8], [17]. Restricting to integer index Sobolev spaces excludes certain interesting cases from study. Most of the results in this paper are proved for the integer case, but

in Section 6, estimates of the form (1.1) involving fractional order Sobolev norms are proved, and an illustration of their application is given in Example 3 of Section 8.

Several of the questions we discuss here have been treated from different points of view by many authors. Our interest in these questions comes from studying the approximation results that are needed to analyze finite element methods. In this area, the work of Bramble and Hilbert [4], [5] is fundamental. The work of Ciarlet and Wagschal on multipoint Taylor formulas [7] is another approach to giving constructive proofs of approximation results needed for finite element analysis and their results played an important role in the evolution of this paper. The basic representation given in Section 3, which we call a Sobolev representation, is quite similar to one used by Sobolev [18] in proving imbedding theorems (see Remark 3.5). However, it appears to be different from the one used in [18] for which, in particular, it is not clear that the commutativity property mentioned above is valid. A more recent treatment of related representations, as well as some discussion of their applications in other work, can be found in an article by Burenkov [6]. In [15], Meinguet gave a constructive polynomial approximation process that is closely related to the Sobolev representation in Section 3 and, in [2], Arcangeli and Gout applied Meinguet's ideas to Lagrange interpolation in  $\mathbf{R}^n$ .

**2. Notation.** Let  $x, y, \dots$  denote points in  $\mathbf{R}^n$ , and let  $dx, dy, \dots$  denote Lebesgue measure. If  $D$  is a measurable set,  $p \in [1, \infty]$ , and  $f$  is a (real or complex valued) measurable function, we say  $f \in L_p(D)$  if

$$(2.1) \quad \|f\|_{L_p(D)} \equiv \left( \int_D |f(x)|^p dx \right)^{1/p} < \infty$$

with the usual modification when  $p = \infty$ . When  $p = \infty$ ,  $1/p$  is defined to be zero.

Let  $\mathbf{N}$  denote the set of nonnegative integers. A *multi-index*  $\alpha$  is an  $n$ -tuple of nonnegative integers:  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $\alpha_i \in \mathbf{N}$ ,  $i = 1, \dots, n$ . We have the following definitions:

$$(2.2) \quad |\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$

$$(2.3) \quad \alpha \leq \beta \quad \text{iff} \quad \alpha_i \leq \beta_i, \quad i = 1, \dots, n,$$

$$(2.4) \quad (\alpha + \beta)_i = \alpha_i + \beta_i, \quad i = 1, \dots, n,$$

$$(2.5) \quad (\alpha - \beta)_i = \max\{\alpha_i - \beta_i, 0\}, \quad i = 1, \dots, n,$$

$$(2.6) \quad \alpha! = (\alpha_1!)(\alpha_2!) \cdots (\alpha_n!),$$

$$(2.7) \quad x^\alpha = (x_1^{\alpha_1})(x_2^{\alpha_2}) \cdots (x_n^{\alpha_n}), \quad \text{and}$$

$$(2.8) \quad \left( \frac{\partial}{\partial x} \right)^\alpha = \left( \frac{\partial}{\partial x_1} \right)^{\alpha_1} \left( \frac{\partial}{\partial x_2} \right)^{\alpha_2} \cdots \left( \frac{\partial}{\partial x_n} \right)^{\alpha_n}.$$

We let  $\delta^i$ ,  $i = 1, \dots, n$ , denote the multi-index whose  $i$ th component is 1 and the

rest are zero:

$$(2.9) \quad \delta_j^i = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

When  $D$  is an open set, denote by  $C^\infty(D)$  the space of infinitely differentiable functions in  $D$ . For  $f \in C^\infty(D)$ , we use the notation

$$(2.10) \quad f^{(\alpha)}(x) = \left( \frac{\partial}{\partial x} \right)^\alpha f(x) \quad (x \in D).$$

Let  $C_0^\infty(D)$  denote the subset of  $C^\infty(D)$  functions that have compact support in  $D$ . Let  $\mathcal{D}(D)$  denote  $C_0^\infty(D)$  topologized with the usual inductive limit topology [13]. The dual  $\mathcal{D}'(D)$  of  $\mathcal{D}(D)$  is called the set of distributions on  $D$ . If  $\phi \in \mathcal{D}'(D)$  and if  $\alpha$  is a multi-index,  $\phi^{(\alpha)}$  is called a distributional or weak derivative of  $\phi$ , where  $\phi^{(\alpha)}$  is defined by

$$\phi^{(\alpha)}(f) = (-1)^{|\alpha|} \phi(f^{(\alpha)}), \quad f \in \mathcal{D}(D).$$

A distribution  $\phi \in \mathcal{D}'(D)$  is identified with a function  $\psi$  defined on  $D$  if for each  $f \in \mathcal{D}(D)$ ,  $\psi f \in L_1(D)$  and  $\phi(f) = \int_D \psi f dx$ . In this case we shall let  $\phi$  denote the identified function,  $\psi$ , as well.

If  $m \in \mathbb{N}$  and if for each  $\alpha \in \mathbb{N}^n$  with  $|\alpha| \leq m$ ,  $\phi^{(\alpha)}$  is given by a function such that

$$(2.11) \quad \|\phi\|_{W_p^m(D)} = \sum_{|\alpha| \leq m} \|\phi^{(\alpha)}\|_{L_p(D)} < \infty,$$

then  $\phi \in W_p^m(D)$ . Note that  $C^\infty(D) \cap W_p^m(D)$  is dense in  $W_p^m(D)$  provided  $p < \infty$ . (See [11, p. 15] for a proof.) If  $D$  has finite measure, then  $W_p^m(D) \subset W_q^m(D)$  if  $1 \leq q \leq p \leq \infty$  (by Hölder's inequality). For  $\phi \in W_p^m(D)$  let

$$(2.12) \quad |\phi|_{W_p^m(D)} = \sum_{|\alpha|=m} \|\phi^{(\alpha)}\|_{L_p(D)}.$$

Let  $r$  be a positive integer, and denote by  $\mathcal{P}_r$  the space of polynomials in  $n$  variables of degree less than  $r$ . Let  $\mathcal{P}_\infty = \bigcup_{r=1}^\infty \mathcal{P}_r$ .

**3. Approximation by Complete Polynomials.** Let  $D$  be a bounded set in  $\mathbb{R}^n$  with diameter  $d$ . Suppose  $D$  is star-shaped with respect to every point in an open ball  $B$ . Let  $\phi \in C_0^\infty(B)$  have integral one. *Throughout Sections 3, 4, 5, and 6,  $D$ ,  $d$ ,  $B$ , and  $\phi$  will remain the same.*

**SOBOLEV'S REPRESENTATION.** If  $f \in C^\infty(D)$ ,  $l$  is a positive integer, and  $x \in D$ , then

$$(3.1) \quad f(x) = Q^l f(x) + R^l f(x),$$

where

$$(3.2) \quad Q^l f(x) = \sum_{|\alpha| \leq l} \int_B \phi(y) f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!} dy$$

is a polynomial of degree less than  $l$  and

$$(3.3) \quad R^l f(x) = \sum_{|\alpha|=l} \int_D k_\alpha(x, y) f^{(\alpha)}(y) dy.$$

The kernels  $k_\alpha$  are given by

$$(3.4) \quad k_\alpha(x, y) = (l/\alpha!) (x - y)^\alpha k(x, y),$$

where

$$(3.5) \quad k(x, y) = \int_0^1 s^{-n-1} \phi(x + s^{-1}(y - x)) ds.$$

*Remark 3.1.* As a function of  $y$ ,  $k(x, \cdot)$ , and, therefore, each  $k_\alpha(x, \cdot)$ , is supported in the convex hull of  $\{x\} \cup \text{supp } \phi$ ; in particular, the region of integration in (3.3) is contained in a compact subset of  $D$ .  $\square$

*Remark 3.2.* Integration by parts shows that  $Q^l$  is defined for all  $f \in \mathcal{D}'(B)$  and that, in particular, for  $f$  in  $L_1(B)$

$$(3.6) \quad \|Q^l f\|_{W_{\infty}^{l-1}(D)} \leq C(n, l, d, \phi) \|f\|_{L_1(B)}. \quad \square$$

*Proof of the Representation.* Let  $x \in D$ ,  $y \in B$ , and use Taylor's theorem:

$$f(x) = \sum_{|\alpha| < l} \frac{(x - y)^\alpha}{\alpha!} f^{(\alpha)}(y) + l \sum_{|\alpha|=l} \frac{(x - y)^\alpha}{\alpha!} \int_0^1 s^{l-1} f^{(\alpha)}(x + s(y - x)) ds.$$

Multiply by  $\phi(y)$  and integrate with respect to  $y$ :

$$f(x) = Q^l f(x) + l \sum_{|\alpha|=l} \frac{1}{\alpha!} \int \phi(y) (x - y)^\alpha \int_0^1 s^{l-1} f^{(\alpha)}(x + s(y - x)) ds dy.$$

Using Fubini's theorem and the change of variables  $z = x + s(y - x)$ , one finds

$$\begin{aligned} & \int \phi(y) (x - y)^\alpha \int_0^1 s^{l-1} f^{(\alpha)}(x + s(y - x)) ds dy \\ &= \int_0^1 \int \phi(y) (x - y)^\alpha s^{l-1} f^{(\alpha)}(x + s(y - x)) dy ds \\ &= \int_0^1 \int \phi(x + s^{-1}(z - x)) (x - z)^\alpha s^{-1} f^{(\alpha)}(z) s^{-n} dz ds \\ &= \int (x - z)^\alpha f^{(\alpha)}(z) \left( \int_0^1 \phi(x + s^{-1}(z - x)) s^{-n-1} ds \right) dz \\ &= \frac{\alpha!}{l} \int k_\alpha(x, z) f^{(\alpha)}(z) dz. \end{aligned}$$

The use of Fubini's theorem is justified because

$$(3.7) \quad \begin{aligned} |k(x, z)| &= \left| \int_0^1 \phi(x + s^{-1}(z - x)) s^{-n-1} ds \right| \\ &= \left| \int_{|z-x|/d}^1 \phi(x + s^{-1}(z - x)) s^{-n-1} ds \right| \leq c_1 |x - z|^{-n}, \end{aligned}$$

where  $c_1 = \|\phi\|_{L_\infty(B)} d^n/n$ .  $\square$

*Remark 3.3.* In view of (3.4), it follows as in (3.7) that

$$(3.8) \quad \left| \left( \frac{\partial}{\partial x} \right)^\beta \left( \frac{\partial}{\partial y} \right)^\gamma k_\alpha(x, y) \right| \leq C(n, l, d, \phi, |\beta|, |\gamma|) |x - y|^{|\alpha| - n - |\beta| - |\gamma|}. \quad \square$$

In view of the Sobolev representation and (3.8), estimates of the approximation error  $u - Q^l u$  may be reduced to consideration of the Riesz potentials

$$(3.9) \quad I^l f(x) \equiv \int_D |x - y|^{l-n} f(y) dy.$$

The following proposition collects several known results in the form we find useful.

**PROPOSITION 3.1.** *Let  $l$  be a positive integer and let  $p$  and  $q$  be in  $[1, \infty]$ . Suppose that  $1/q - 1/p + l/n \geq 0$  and that  $\sigma$  is a positive lower bound for*

$$\max \left\{ \left\lceil \frac{l}{n} \right\rceil, \frac{1}{q} - \frac{1}{p} + \frac{l}{n}, \min \left\{ 1 - \frac{1}{p}, \frac{1}{q} \right\} \right\},$$

where  $[x]$  is the largest integer not greater than  $x$ . Then  $I^l$  maps  $L_p(D)$  to  $L_q(D)$  and for all  $f \in L_p(D)$

$$\|I^l f\|_{L_q(D)} \leq C(n, l, d, \sigma) \|f\|_{L_p(D)}.$$

*Proof.* Let

$$R_l(x) = \begin{cases} |x|^{l-n} & \text{for } x \leq d, \\ 0 & \text{for } x > d. \end{cases}$$

Then, for  $x \in D$ ,

$$I^l f(x) = R_l * f(\tilde{x}),$$

where  $f$  is extended to  $\mathbf{R}^n$  by zero outside  $D$ .

For  $l \geq n$ ,  $R_l$  is bounded and, hence,

$$\|I^l f\|_{L_\infty(D)} \leq C(n, l, d) \|f\|_{L_1(D)}.$$

The result then follows from Hölder's inequality.

Now suppose that  $l < n$  and that  $1/q - 1/p + l/n \geq \sigma > 0$ . For this case we apply Young's inequality [21] to obtain

$$\|I^l f\|_{L_q(D)} = \|R_l * f\|_{L_q(D)} \leq \|R_l\|_{L_r(\mathbf{R}^n)} \|f\|_{L_p(D)},$$

where  $1/r = 1 - 1/p + 1/q \geq 1 - l/n + \sigma$ . But

$$\|R_l\|_{L_r(\mathbf{R}^n)} = C(n) \left[ \frac{d^{(l-n)r+n}}{(l-n)r+n} \right]^{1/r} < \infty,$$

since  $(l-n)r+n \geq \sigma nr$ . Hence, the proposition holds in this case as well.

Now suppose that  $l < n$  and  $1/q - 1/p + l/n = 0$ . Then it is a standard result (cf. Stein [19, p. 119]) that

$$\|I^l f\|_{L_q(\mathbf{R}^n)} \leq C(n, l, p, q) \|f\|_{L_p(\mathbf{R}^n)},$$

provided that  $p > 1$  and  $p < n/l$ , i.e.,  $q < \infty$ . It is clear from the proof in Stein referenced above that the constant  $C(n, l, p, q)$  can be chosen to be continuous in  $p$  and  $q$ , and hence bounded by

$$C(n, l, \sigma) = \sup \left\{ C(u, l, p, q) : \max \left( 1 - \frac{1}{p}, \frac{1}{q} \right) \geq \sigma \right\} < \infty.$$

This yields the proposition in the case  $1/q - 1/p + l/n = 0$  and  $\max(1 - 1/p, 1/q) \geq \sigma > 0$ .

Now suppose  $1/q - 1/p + l/n$  is positive, but arbitrarily small, and  $\max(1 - 1/p, 1/q) \geq \sigma > 0$ . If  $1 - 1/p \geq 1/q$ , choose  $\tilde{q}$  such that  $1/\tilde{q} - 1/p + l/n = 0$ . Since  $\max(1 - 1/p, 1/\tilde{q}) \geq \sigma$ , the previous case implies that

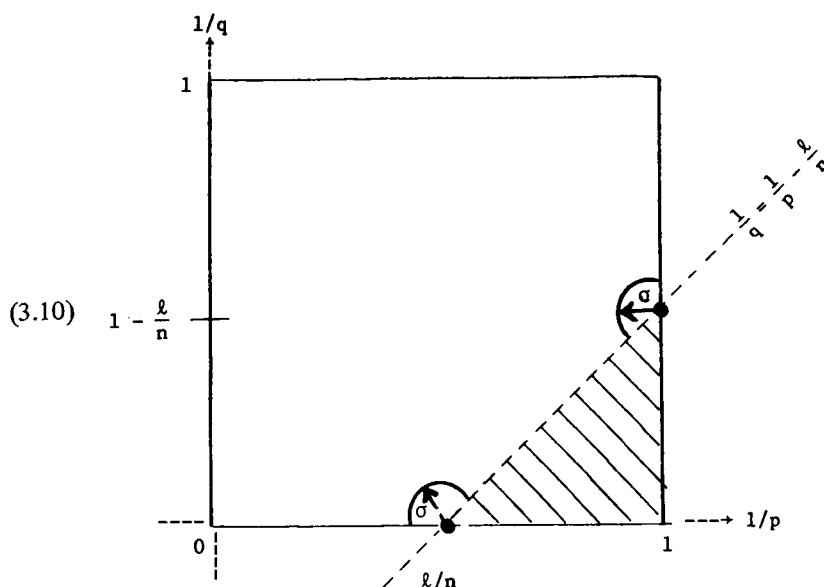
$$\|I^l f\|_{L_{\tilde{q}}(D)} \leq C(n, l, d, \sigma) \|f\|_{L_p(D)}$$

for all  $f \in L_p(D)$ . Since  $1/\tilde{q} < 1/q$ , Hölder's inequality yields the desired result. If on the other hand  $1 - 1/p \leq 1/q$ , choose  $p$  such that  $1/q - 1/\tilde{p} + l/n = 0$ . The previous case again implies that

$$\|I^l f\|_{L_q(D)} \leq C(n, l, d, \sigma) \|f\|_{L_{\tilde{p}}(D)},$$

and Hölder's inequality applied to the right-hand side yields the desired conclusion. This completes the proof of the proposition.  $\square$

*Remark 3.4.* The above proposition may be interpreted via the following  $1/q$  v.s.  $1/p$  diagram:



The proposition holds for all pairs  $(1/p, 1/q)$  in the closed unit square excluding the shaded region lying below the line  $1/q = 1/p - l/n$  and excluding the two points  $(l/n, 0)$  and  $(1, l/n)$ . Furthermore, the norm of  $I^l: L_p(D) \rightarrow L_q(D)$  can be bounded uniformly in the closed subset of the unit square excluding the shaded region and excluding discs of radius  $\sigma$  around the points  $(l/n, 0)$  and  $(1, l/n)$ . However, as  $\sigma$  is

allowed to tend to zero, the norm of  $I^l$  tends to infinity. (If  $l \geq n$ , then  $I$  is bounded uniformly for all  $p$  and  $q$ .) Notice that for all  $l, n, p$ , and  $q$  for which the proposition is applicable, it is also applicable for  $l, n, p'$ , and  $q$  for some  $p' < \infty$ . This observation will be used later to restrict attention to finite  $p$  in order to allow the use of a density argument.

The restriction that  $p \neq 1$  and  $q \neq \infty$  when  $1/q = 1/p - l/n$  is necessary since the Riesz potential of order  $l$  does not map  $L_1(\mathbb{R}^n)$  (respectively,  $L_{n/l}(\mathbb{R}^n)$ ) into  $L_{n/n-l}(\mathbb{R}^n)$  (respectively,  $L_\infty(\mathbb{R}^n)$ ); see Stein [19, p. 119]. However that the case  $p = 1$  may be treatable by another argument is indicated by the fact that the Sobolev embedding holds in this case; see Stein [19, pp. 128–129].  $\square$

Viewing the Sobolev representation of  $f$  as giving a polynomial approximation  $(Q^l f)$  to  $f$ , there are now two natural polynomial approximations to the derivatives of  $f$ , namely,  $(\partial/\partial x)^\alpha Q^l f$  and  $Q^{l-|\alpha|}(\partial/\partial x)^\alpha f$ . Both are polynomials of degree less than  $l - |\alpha|$ . Schematically,

$$(3.11) \quad \begin{array}{ccc} \mathcal{D} & \xrightarrow{\left(\frac{\partial}{\partial x}\right)^\alpha} & \mathcal{D}' \\ \downarrow Q^l & & \downarrow Q^{l-|\alpha|} \\ \mathcal{P}_l & \xrightarrow{\left(\frac{\partial}{\partial x}\right)^\alpha} & \mathcal{P}_{l-|\alpha|} \end{array}$$

**THEOREM 3.1.** *The diagram (3.11) commutes, i.e. for  $f \in \mathcal{D}'(B)$ ,  $(\partial/\partial x)^\alpha Q^l f = Q^{l-|\alpha|}(\partial/\partial x)^\alpha f$ .*

*Proof.* Let  $f \in C^\infty(B)$  and let  $x, y \in B$ . We write the Taylor polynomial of  $f$  as

$$T_y^l f(x) = \sum_{|\alpha| \leq l} f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!}.$$

Then

$$\left(\frac{\partial}{\partial x}\right)^\beta (T_y^l f)(x) = T_y^{l-|\beta|} \left(\left(\frac{\partial}{\partial x}\right)^\beta f\right)(x),$$

as is easily proved by induction. But since

$$Q^l f(x) = \int \phi(y) T_y^l f(x) dy,$$

the result follows by differentiation under the integral.

The result then follows for  $f \in \mathcal{D}'(B)$  since  $(\partial/\partial x)^\alpha$  is continuous on  $\mathcal{D}'(B)$ ,  $Q^k$  is continuous from  $\mathcal{D}'(B)$  into  $\mathcal{P}_k$ , and  $C^\infty(B)$  is dense in  $\mathcal{D}'(B)$  [13, p. 15].  $\square$

Combining Proposition 3.1 with Theorem 3.1 gives the following:

**THEOREM 3.2.** *Let  $m$  and  $l$  be integers such that  $0 \leq m < l$ , and let  $p$  and  $q$  be in  $[1, \infty]$ . Suppose that, with  $\tilde{l} = l - m$ ,  $1/q - 1/p + \tilde{l}/n \geq 0$  and that*

$$0 < \sigma \leq \max \left\{ \lceil \tilde{l}/n \rceil, \frac{1}{q} - \frac{1}{p} + \frac{\tilde{l}}{n}, \min \left\{ 1 - \frac{1}{p}, \frac{1}{q} \right\} \right\}.$$



Then for  $f \in W_p^l(D)$

$$\|f - Q^l f\|_{W_q^m(D)} \leq C(n, l, d, \phi, \sigma) |f|_{W_p^l(D)}.$$

*Proof.* In view of Remark 3.4, it suffices to assume that  $p < \infty$ , for then the general case follows from Hölder's inequality. Since  $C^\infty(D) \cap W_p^l(D)$  is dense in  $W_p^l(D)$ , it suffices to prove the estimate for  $f \in C^\infty(D) \cap W_p^l(D)$ . Take  $\alpha \in \mathbb{N}^n$  such that  $|\alpha| \leq m$ . Then by Theorem 1,

$$\left(\frac{\partial}{\partial x}\right)^\alpha (f - Q^l f) = f^{(\alpha)} - Q^{l-|\alpha|} f^{(\alpha)} = R^{l-|\alpha|} f^{(\alpha)}.$$

Thus the result follows from (3.8) and Proposition 3.1.  $\square$

*Remark 3.5.* Sobolev's proof [18] of the imbedding theorems was essentially via Theorem 3.2, Remark 3.2, and the triangle inequality:

$$\begin{aligned} \|f\|_{W_q^m(D)} &\leq \|f - Q^l f\|_{W_q^m(D)} + \|Q^l f\|_{W_q^m(D)} \\ &\leq C_1 |f|_{W_p^l(D)} + C_2 \|f\|_{L_1(D)}. \end{aligned}$$

It is not clear that Theorem 3.1 holds for the representation used by Sobolev, but as noted below, Theorem 3.2 does not really rely on the commutativity.  $\square$

*Remark 3.6.* The estimate in Theorem 3.2 could likewise be derived without using Theorem 3.1 simply by differentiating under the integral, in view of (3.8). However, the use of the commutativity becomes crucial in the next two sections.  $\square$

*Remark 3.7.* Note that if  $m = l$ , then the conclusion of Theorem 3.2 remains valid for  $q \leq p$ . This follows because

$$|f - Q^l f|_{W_q^l(D)} = |f|_{W_q^l(D)}. \quad \square$$

**4. Extended Tensor-Product Polynomial Approximation.** Let  $A$  be a set of multi-indices, and let the polar of  $A$ ,  $A^0$ , be the set of multi-indices given by

$$(4.1) \quad A^0 = \left\{ \beta \in \mathbb{N}^n : \left(\frac{\partial}{\partial x}\right)^\alpha x^\beta \equiv 0 \text{ for all } \alpha \in A \right\}.$$

If  $A$  and  $B$  are two sets of multi-indices such that  $A \supset B$ , then  $A^0 \subset B^0$ .

Two sets of multi-indices that play important roles are the following:

$$A = \{\alpha \in \mathbb{N}^n : |\alpha| = l\}, \quad B = \{l\delta^1, \dots, l\delta^n\}.$$

In these cases

$$A^0 = \{\beta : |\beta| < l\}, \quad B^0 = \{\beta : \beta_i < l \text{ for } i = 1, \dots, n\}.$$

The set  $A^0$  is naturally associated with complete polynomials of degree less than  $l$  while  $B^0$  is naturally associated with polynomials that are of degree less than  $l$  in each variable separately.

For any set of multi-indices  $A$  define the base of  $A$ ,  $A_-$ , as the collection of all  $\alpha \in A$  such that  $\beta \in A$  and  $\beta \leq \alpha$  implies that  $\beta = \alpha$ . Note that  $A^0 = (A_-)^0$  since

$\alpha \in A$  implies that there is  $\gamma \in A_-$  such that  $\gamma \leq \alpha$ , and hence  $(\partial/\partial x)^\gamma x^\beta \equiv 0$  implies  $(\partial/\partial x)^\alpha x^\beta \equiv 0$ .

LEMMA 4.1.  $A^0$  is a finite set if and only if there are nonnegative integers  $r_i$ ,  $i = 1, \dots, n$ , such that

$$\{r_i \delta^i: i = 1, \dots, n\} \subset A.$$

*Proof.* The “if” is obvious, since

$$A^0 \subset \{r_i \delta^i: i = 1, \dots, n\}^0 = \{\alpha: \alpha_i < r_i, i = 1, \dots, n\}.$$

To prove “only if”, suppose that, for some  $i$ ,  $\{n \delta^i: n \in \mathbb{N}\} \cap A = \emptyset$ . Then for  $\alpha \in A$ , there is some  $j \neq i$  such that  $\alpha_j \neq 0$ , and so if  $f$  is any function that is constant as a function of  $x_j$ , then  $f^{(\alpha)} \equiv 0$ . In particular, if  $f$  depends on  $x_i$  alone, then  $f^{(\alpha)} \equiv 0$  for all  $\alpha \in A$ . Thus,  $\{n \delta^i: n \in \mathbb{N}\} \subset A^0$ , and hence,  $A^0$  is not finite.  $\square$

Remark 4.1. If  $A^0$  is a finite set, then it follows from Lemma 4.1 that

$$\max_{\beta \in A^0} |\beta| < n \left( \max_{\alpha \in A_-} |\alpha| - 1 \right).$$

Note that it also follows from Lemma 4.1 that if  $A^0$  is finite then the class of polynomials spanned by  $x^\beta$  for  $\beta \in A^0$  is a subset of the tensor product space of polynomials which are of degree less than  $r_i$  in  $x_i$  for  $i = 1, \dots, n$ .  $\square$

EXTENDED TENSOR PRODUCT REPRESENTATION. Given a finite set  $A$  of multi-indices such that  $A^0$  is finite and given  $f \in C^\infty(D)$ ,

$$(4.2) \quad f(x) = Q^A f(x) + R^A f(x), \quad x \in D,$$

where

$$(4.3) \quad Q^A f(x) = \sum_{\alpha \in A^0} \int_B \phi(y) f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!} dy$$

and

$$(4.4) \quad R^A f(x) = \sum_{\alpha \in A_-} \int_D \bar{k}_\alpha(x, y) f^{(\alpha)}(y) dy.$$

If  $A^0 = \emptyset$ , then the sum over  $A^0$  is identically zero. The kernels  $\bar{k}_\alpha$  satisfy

$$(4.5) \quad \left| \left( \frac{\partial}{\partial x} \right)^\beta \left( \frac{\partial}{\partial y} \right)^\gamma \bar{k}_\alpha(x, y) \right| \leq C(l, n, \phi, d, |\beta|, |\gamma|) |x-y|^{|\alpha|-n-|\beta|-|\gamma|},$$

where  $l = 1 + \max_{\alpha \in A^0} |\alpha|$ .

*Proof.* Consider Sobolev's representation of order  $l$ :

$$\begin{aligned} f(x) &= Q^l f(x) + R^l f(x) \\ &= \sum_{|\alpha| < l} \int \phi(y) f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!} dy + \sum_{|\alpha|=l} \int k_\alpha(x, y) f^{(\alpha)}(y) dy. \end{aligned}$$

The set of all multi-indices  $\mathbf{N}^n$  decomposes into two disjoint sets, namely  $A^0$  and  $\bar{A} \equiv \{\beta: \exists \alpha \in A \ni \alpha \leq \beta\} = A + \mathbf{N}^n = A_- + \mathbf{N}^n$ . Since  $l > (\max_{\beta \in A^0} |\beta|)$ ,  $|\alpha| = l$  implies that  $\alpha \in \bar{A}$ . Thus, we have

$$f(x) = Q^A f(x) + \sum_{\substack{|\alpha| < l \\ \alpha \notin A^0}} \int \phi(y) f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!} dy + \sum_{|\alpha|=l} \int k_\alpha(x, y) f^{(\alpha)}(y) dy,$$

and the sums in the remainder terms are over  $\alpha \in \bar{A}$ . It remains to convert these terms to the form (4.4). But, for each  $\alpha \in \bar{A}$ , there is some  $\beta \in A_-$  such that  $\beta \leq \alpha$ , so we may write  $f^{(\alpha)} = (\partial/\partial y)^{\alpha-\beta} f^{(\beta)}$ , integrate by parts  $|\alpha - \beta|$  times, and obtain

$$\begin{aligned} \int \phi(y) f^{(\alpha)}(y) \frac{(x-y)^\alpha}{\alpha!} dy \\ = (-1)^{|\alpha-\beta|} \int \left\{ \left( \frac{\partial}{\partial y} \right)^{\alpha-\beta} \left[ \phi(y) \frac{(x-y)^\alpha}{\alpha!} \right] \right\} f^{(\beta)}(y) dy \end{aligned}$$

or

$$\int k_\alpha(x, y) f^{(\alpha)}(y) dy = (-1)^{|\alpha-\beta|} \int \left\{ \left( \frac{\partial}{\partial y} \right)^{\alpha-\beta} k_\alpha(x, y) \right\} f^{(\beta)}(y) dy.$$

Summing over all  $\alpha$ , we obtain (4.4); it is not clear whether the  $\bar{k}_\beta$ 's are uniquely determined by the above process. Estimate (4.5) follows from (3.8).  $\square$

We now consider the commutativity of  $Q^A$  with differentiation. For two multi-indices  $\alpha$  and  $\beta$ , (2.5) defines a new multi-index  $\alpha - \beta$ . Note that  $\alpha - \beta$  is defined even if  $\alpha \not\geq \beta$  and that  $(\alpha - \beta) + \beta \geq \alpha$ , with equality if and only if  $\alpha \geq \beta$ . Given a set  $A$  of multi-indices and a multi-index  $\beta$ , define a new set  $A - \beta \subset \mathbf{N}^n$  by

$$A - \beta = \{\alpha - \beta: \alpha \in A\}.$$

Since  $(A - \beta)^0 \subset A^0$ , we see that if  $A^0$  is finite then so is  $(A - \beta)^0$ .

**THEOREM 4.1.** *Let  $A$  be a finite set of multi-indices such that  $A^0$  is finite and let  $\beta$  be a multi-index. Then*

$$(4.6) \quad \left( \frac{\partial}{\partial x} \right)^\beta Q^A f = Q^{A-\beta} f^{(\beta)} \quad \text{for any } f \in \mathcal{D}'(B).$$

*Proof.* It is easy to see that  $(\partial/\partial x)^{\gamma-\beta} x^\delta \equiv 0$  if and only if  $(\partial/\partial x)^\gamma x^{\delta+\beta} \equiv 0$ . Thus,  $\delta \in (A - \beta)^0$  if and only if  $\delta + \beta \in A^0$ . Hence,

$$(A - \beta)^0 = \{\delta - \beta: \delta \in A^0, \delta \geq \beta\}.$$

Thus, for  $f \in C^\infty(B)$ ,

$$\begin{aligned} \left( \frac{\partial}{\partial x} \right)^\beta Q^A f(x) &= \sum_{\substack{\gamma \in A^0 \\ \gamma \geq \beta}} \int \phi(y) f^{(\gamma)}(y) \frac{(x-y)^{\gamma-\beta}}{(\gamma-\beta)!} dy \\ &= \sum_{\delta \in (A-\beta)^0} \int \phi(y) f^{(\delta+\beta)}(y) \frac{(x-y)^\delta}{\delta!} dy = Q^{A-\beta} f^{(\beta)}(x). \end{aligned}$$

The result follows for  $f \in \mathcal{D}'(B)$  by density.  $\square$

*Remark 4.2.* Note that both sides in (4.6) are zero unless  $\beta \in A^0$ .  $\square$

The remaining results in this section and the results in Section 5 hold for functions that are in a function space described in Remark 5.4. In each case it suffices to prove these results for functions in  $C^\infty(D)$  such that all relevant norms are finite. The definition of the function space is delayed so that it need be given only once and because the intervening results make the appropriateness of the norm used much more apparent.

**THEOREM 4.2.** *Let  $A$  be a finite set of multi-indices such that  $A^0$  is a finite set; let  $\beta$  be any multi-index; let  $l = \min_{\alpha \in A - \beta} |\alpha|$ ; let  $m$  be a nonnegative integer less than  $l$ ; and let  $\{q\} \cup \{p_\alpha: \alpha \in (A - \beta)_-\} \subset [1, \infty]$ . Suppose that*

$$\min \left\{ \frac{1}{q} - \frac{1}{p_\alpha} + (|\alpha| - m)/n: \alpha \in (A - \beta)_- \right\} \geq 0,$$

*and that  $0 < \sigma \leq \min\{\mu_\alpha: \alpha \in (A - \beta)_-\}$ , where, with  $\tilde{\alpha} = |\alpha| - m$ ,*

$$\mu_\alpha = \max \left\{ [\tilde{\alpha}/n], \frac{1}{q} - \frac{1}{p_\alpha} + \frac{\tilde{\alpha}}{n}, \min \left\{ 1 - \frac{1}{p_\alpha}, \frac{1}{q} \right\} \right\}.$$

*Then*

$$\left\| \left( \frac{\partial}{\partial x} \right)^\beta (f - Q^A f) \right\|_{W_q^m(D)} \leq C(n, A, \beta, m, d, \phi, \sigma) \sum_{\alpha \in (A - \beta)_-} \|f^{(\alpha + \beta)}\|_{L_{p_\alpha}(D)}.$$

*Proof.* Use Theorem 4.1 to see that

$$\left( \frac{\partial}{\partial x} \right)^\beta (f - Q^A f) = R^{A - \beta} f^{(\beta)}.$$

Thus, it suffices to prove the result for  $\beta = 0$ . Differentiating under the integral in (4.4) and using (4.5) and Proposition 3.1 completes the proof.  $\square$

## 5. Generalized Polynomial Approximation.

**THEOREM 5.1.** *Let  $P_1, \dots, P_k$  be nontrivial homogeneous polynomials (in  $n$  variables) of degrees  $l_1, \dots, l_k$ , respectively, having no common (nonzero) complex zero (this forces  $k \geq n$ ). Define*

$$K = \left\{ f \in \mathcal{D}'(\mathbb{R}^n): P_j \left( \frac{\partial}{\partial x} \right) f \equiv 0 \text{ for } j = 1, \dots, k \right\}.$$

*Then  $K \subset \mathcal{P}_r$  for some integer  $r$ . Let  $l = \min_{1 \leq j \leq k} l_j$ ; let  $m$  be a nonnegative integer less than  $l$ ; and let  $\{p_j: j = 1, \dots, k\} \subset [1, \infty]$ . Suppose that*

$$\frac{1}{q} - \frac{1}{p_j} + (l_j - m)/n \geq 0, \quad j = 1, \dots, k,$$

*and that  $0 < \sigma \leq \min\{\mu_j: j = 1, \dots, k\}$  where, with  $\tilde{l}_j = l_j - m$ ,*

$$\mu_j = \max \left\{ [\tilde{l}_j/n], \frac{1}{q} - \frac{1}{p_j} + \tilde{l}_j/n, \min \left\{ 1 - \frac{1}{p_j}, \frac{1}{q} \right\} \right\}.$$

*Then*

$$\inf_{Q \in K} \|f - Q\|_{W_q^m(D)} \leq C(n, m, \{P_j\}, d, \phi, \sigma) \sum_{j=1}^k \left\| P_j \left( \frac{\partial}{\partial x} \right) f \right\|_{L_{p_j}(D)}.$$

*Proof.* As is Agmon [1], it follows from Hilbert's Nullstellensatz that there is an integer  $r$  such that for all  $|\alpha| = r$ ,

$$(5.1) \quad \xi^\alpha = \sum_{j=1}^k R_j^\alpha(\xi) P_j(\xi)$$

for some polynomials  $R_j^\alpha$  that are homogeneous of degree  $r - l_j$ . Thus each  $f \in K$  satisfies  $f^{(\alpha)} \equiv 0$  for all  $|\alpha| = r$ , i.e.,  $K \subset P_r$ .

Since  $K \subset P_r$ , it follows that for any  $P \in P_r$

$$\inf_{Q \in K} \|P - Q\|_{W_q^m(D^*)} \leq C \sum_{j=1}^k \left\| P_j \left( \frac{\partial}{\partial x} \right) P \right\|_{L_{p_j}(B)},$$

where  $D^*$  is the ball of diameter  $2d$  concentric with  $B$ , because of the equivalence of norms on the finite-dimensional space  $P_r/K$ . Therefore,

$$(5.2) \quad \inf_{Q \in K} \|P - Q\|_{W_q^m(D)} \leq C \sum_{j=1}^k \left\| P_j \left( \frac{\partial}{\partial x} \right) P \right\|_{L_{p_j}(D)},$$

with  $C$  depending only on the diameter  $d$  of  $D$ , the diameter of  $B$ , and  $\{P_j\}$ . (The independence from  $q$  and  $\{p_j\}$  can be achieved using Hölder's inequality.) Using Sobolev's representation of order  $r$  and the triangle inequality, we get

$$\begin{aligned} \inf_{Q \in K} \|f - Q\|_{W_q^m(D)} &\leq \|f - Q^r f\|_{W_q^m(D)} + \inf_{Q \in K} \|Q^r f - Q\|_{W_q^m(D)} \\ &\leq \|R^r f\|_{W_q^m(D)} + C \sum_{j=1}^k \left\| P_j \left( \frac{\partial}{\partial x} \right) Q^r f \right\|_{W_q^m(D)}. \end{aligned}$$

Because of Theorem 3.1 and the linearity of  $Q^r$ ,

$$P_j \left( \frac{\partial}{\partial x} \right) Q^r f = \left[ \sum_{|\alpha|=l_j} c_{\alpha,j} \left( \frac{\partial}{\partial x} \right)^\alpha \right] Q^r f = \sum_{|\alpha|=l_j} c_{\alpha,j} Q^{r-l_j} f^{(\alpha)} = Q^{r-l_j} P_j \left( \frac{\partial}{\partial x} \right) f.$$

Thus, Remark 3.2 shows that

$$\left\| P_j \left( \frac{\partial}{\partial x} \right) Q^r f \right\|_{W_q^m(D)} \leq C \left\| P_j \left( \frac{\partial}{\partial x} \right) f \right\|_{L_1(D)} \leq C \left\| P_j \left( \frac{\partial}{\partial x} \right) f \right\|_{L_{p_j}(D)}.$$

It now remains to estimate  $R^r f$ . Using (5.1), we have

$$\begin{aligned} R^r f(x) &= \sum_{|\alpha|=r} \int k_\alpha(x, y) f^{(\alpha)}(y) dy \\ &= \sum_{|\alpha|=r} \sum_{j=1}^k \int k_\alpha(x, y) \left[ R_j^\alpha \left( \frac{\partial}{\partial y} \right) P_j \left( \frac{\partial}{\partial y} \right) f \right](y) dy \\ &= \sum_{j=1}^k \int \left[ \sum_{|\alpha|=r} (-1)^{\deg R_j^\alpha} R_j^\alpha \left( \frac{\partial}{\partial y} \right) k_\alpha(x, y) \right] P_j \left( \frac{\partial}{\partial y} \right) f(y) dy. \end{aligned}$$

Since  $\deg R_j^\alpha = r - l_j$ , (3.8) and Proposition 3.1 imply that

$$\|R^r f\|_{w_q^m(D)} \leq C \sum_{j=1}^k \left\| P_j \left( \frac{\partial}{\partial x} \right) f \right\|_{L_{p_j}(D)},$$

and this completes the proof of the theorem.  $\square$

**Remark 5.1.** If  $A = \{\alpha^1, \dots, \alpha^k\}$  is a finite set of multi-indices and  $\{P_j\}$  is defined by  $P_j(x) = x^{\alpha^j}$  for  $j = 1, \dots, k$ , then  $A^0$  is a finite set if and only if  $P_1, \dots, P_k$  have no common (nonzero) complex zero. This is because  $K \cap P_\infty$  is the space spanned by  $\{x^\beta: \beta \in A^0\}$ ; hence,  $\dim K \cap P_\infty = \text{card } A^0$ . The proof of Theorem 5.1 thus contains the “if” part of our assertion. To prove the “only if” part, suppose there is some  $\xi \neq 0$  such that  $P_j(\xi) = \xi^{\alpha^j} = 0$  for  $j = 1, \dots, k$ . Then some component, say  $\xi_i$ , of  $\xi$  must be nonzero, and so none of the  $\alpha^j$ 's can be of the form  $r\delta^i$ ,  $r \in \mathbb{N}$ ; hence,  $A^0$  is not finite (Lemma 4.1).  $\square$

**Remark 5.2.** The proof of Theorem 5.1 is constructive to the extent that the constant  $C$  in (5.2) can in principle be computed (it is a finite-dimensional problem). The integer  $r$  guaranteed by Hilbert's Nullstellensatz depends only on  $l_1, \dots, l_k$ ; cf. van der Waerden [20, v. 2, p. 6].  $\square$

**Remark 5.3.** If  $0 \neq \xi \in C^n$  is such that  $P_j(\xi) = 0$  for  $j = 1, \dots, k$ , then  $P_j(\partial/\partial x)e^{\xi x} = P_j(\xi)e^{\xi x} = 0$ , so that  $K \not\subset P_\infty$ . Thus, this condition is necessary for polynomial approximation theory. (Note that even if the  $P_j$ 's have real coefficients, there is a real-valued nonpolynomial function in  $K$ , namely,

$$P_j(\partial/\partial x)(\text{Re } e^{\xi x}) = \text{Re } (P_j(\partial/\partial x)e^{\xi x}) = 0,$$

so it is necessary to consider all complex zeros of the  $P_j$ 's.)  $\square$

**Remark 5.4.** The estimates of Sections 4 and 5 are valid for functions  $f$  in the space  $H$  defined as follows: Let  $\{P_j\}$  and  $\{p_j\}$  be finite sets of polynomials and extended real numbers as in Theorem 5.1. Take  $H$  to be the subset of  $L_1(D)$  consisting of functions  $f$  such that the distributional derivatives  $P_j(\partial/\partial x)f$  are elements of  $L_{p_j}(D)$ . This is a Banach space with the norm

$$\|f\|_{L_1(D)} + \sum_j \left\| P_j \left( \frac{\partial}{\partial x} \right) f \right\|_{L_{p_j}(D)}.$$

Further, when all the  $p_j$ 's are finite, the set  $C^\infty(D) \cap H$  is dense in  $H$  (see below), and this allows the proofs of Section 3 to be carried through in view of Remark 3.4.

The claimed density of  $C^\infty(D) \cap H$  is not easily seen by the standard partition of unity argument (cf. [11]), but can be demonstrated as follows: Assume that  $0 \in B$ , and for  $r \geq 1$  let  $f_r(x) = f(x/r)$  and  $D_r = \{x: x/r \in D\}$ . Since  $D$  is star-shaped with respect to  $B$ ,  $D \subset D_r$  for  $r > 1$ . Given  $f \in H$ , it is easily seen that  $f_r \in H$  and that  $f_r \rightarrow f$  in  $H$  as  $r \downarrow 1$ . Thus, it suffices to approximate  $f_r$  by a function in  $C^\infty$ . If  $\psi \in C_0^\infty(\mathbb{R}^n)$ ,  $\int \psi(x)dx = 1$ , and  $\psi_\epsilon(x) = \epsilon^{-n}\psi(x/\epsilon)$ , then as  $\epsilon \downarrow 0$   $\psi_\epsilon * f_r \rightarrow f_r$  in  $L_p$  of any compact subset of  $D_r$  provided  $p < \infty$ . Hence,  $\psi_\epsilon * f_r \rightarrow f_r$  in  $L_1(D)$ . Finally, since  $P_j(\partial/\partial x)(\psi_\epsilon * f_r) = \psi_\epsilon * P_j(\partial/\partial x)f_r$ , we see that  $\psi_\epsilon * f_r \rightarrow f_r$  in  $H$  as  $\epsilon \downarrow 0$ .  $\square$

**6. Approximation by Complete Polynomials of Functions in Fractional Order Sobolev Spaces.** For  $m > 0$  let  $m = \bar{m} + \theta$ , where  $\bar{m}$  is the integer part of  $m$ ; i.e.,  $\bar{m} \in \mathbb{N}$  and  $0 \leq \theta < 1$ . For positive nonintegral  $m$  and  $1 \leq p < \infty$  define

$$(6.1a) \quad |f|_{W_p^m(D)}^p = \sum_{|\alpha|=\bar{m}} \iint_{D \times D} \frac{|f^{(\alpha)}(x) - f^{(\alpha)}(y)|^p}{|x - y|^{n+\theta p}} dx dy.$$

For  $p = \infty$  define the seminorm by

$$(6.1b) \quad |f|_{W_\infty^m(D)} = \sum_{|\alpha|=\bar{m}} \operatorname{ess\,sup}_{D \times D} \frac{|f^{(\alpha)}(x) - f^{(\alpha)}(y)|}{|x - y|^\theta}.$$

The space  $W_p^m(D)$  is the set of all  $\bar{W}_p^m(D)$  functions such that  $|f|_{W_p^m(D)} < \infty$ , and its norm is defined by

$$\|f\|_{W_p^m(D)} = \|f\|_{\bar{W}_p^m(D)} + |f|_{W_p^m(D)}.$$

**PROPOSITION 6.1.** Suppose that  $1 \leq p \leq \infty$ , that  $m = \bar{m} + \theta$  where  $\bar{m} \in \mathbb{N}$  and  $0 < \theta < 1$ , and that  $l = \bar{m} + 1$ . Then there is a constant  $C = C(n, \phi, d, m)$  such that for  $f \in W_p^m(D)$

$$(6.2) \quad \|f - Q^l f\|_{L_p(D)} \leq C |f|_{W_p^m(D)},$$

where  $Q^l$  is defined in (3.2).

*Proof.* First take  $1 \leq p < \infty$ . Then, as in Section 3 we can assume without loss of generality that  $f \in C^\infty(\mathbb{R}^n)$ . (See Grisvard [12].)

Suppose that  $\alpha$  is a multi-index such that  $|\alpha| = l$ , and take  $j$  to be such that  $\alpha = \beta + \delta^j$  where  $\beta$  is a multi-index. Let

$$(6.3) \quad R_\alpha(x) = \int_D f^{(\alpha)}(y) k_\alpha(x, y) dy,$$

where  $k_\alpha$  is defined in (3.4). This can be written as follows:

$$(6.4) \quad \begin{aligned} R_\alpha(x) &= \int_D \frac{\partial}{\partial y_j} [f^{(\beta)}(y) - f^{(\beta)}(x)] k_\alpha(x, y) dy \\ &= \lim_{\epsilon \searrow 0} \left\{ - \int_{\{y \in D: |x-y| > \epsilon\}} [f^{(\beta)}(y) - f^{(\beta)}(x)] \frac{\partial}{\partial y_j} k_\alpha(x, y) dy \right. \\ &\quad \left. + \int_{|x-y|=\epsilon} [f^{(\beta)}(y) - f^{(\beta)}(x)] k_\alpha(x, y) (x_j - y_j) \epsilon^{-1} ds \right\}, \end{aligned}$$

where  $ds$  is surface measure.

The surface integral in (6.4) tends to zero as  $\epsilon \searrow 0$ . To see this note that, for  $|x - y| = \epsilon$ , (3.8) implies

$$|k_\alpha(x, y)| \leq C \epsilon^{1-n},$$

that, for  $|x - y| = \epsilon$ ,  $f \in C^\infty(\mathbb{R}^n)$  implies

$$|f^{(\beta)}(y) - f^{(\beta)}(x)| \leq C \epsilon,$$

and that

$$\int_{|x-y|=\epsilon} 1 \, ds = C\epsilon^{n-1}.$$

Using (3.8) again, we see that

$$(6.5) \quad |R_\alpha(x)| \leq C_0 \int_D \frac{|f^{(\beta)}(y) - f^{(\beta)}(x)|}{|x-y|^n} dy.$$

Note that the integrand is in  $L^1$  since it is bounded by  $C|x-y|^{-n+1}$ . Hölder's inequality and (6.5) imply that

$$(6.6) \quad |R_\alpha(x)|^p \leq C \int_D \frac{|f^{(\beta)}(y) - f^{(\beta)}(x)|^p}{|x-y|^{n+\theta p}} dy.$$

Integrating this with respect to  $x$  and summing on  $|\alpha| = l$  gives (6.2) for the case  $p < \infty$ .

Note that the  $C$  in (6.6) is just

$$C_0^p \left( \int_D |x-y|^{-n+\theta p'} dy \right)^{p/p'} \leq \left\{ C_0 \left( \frac{\omega_n}{\theta} \left( 1 - \frac{1}{p} \right) \right)^{1-1/p} d^\theta \right\}^p,$$

where  $\omega_n$  is the measure of the unit  $(n-1)$ -sphere. Thus the constant  $C$  in (6.2) can be taken to be independent of  $p \in (1, \infty)$ , and it is bounded for  $\theta$  in the interval  $(\epsilon, 1)$  where  $\epsilon$  is positive.

The estimate for  $p = \infty$  is complicated by the facts that, for nonintegral  $m$ ,  $W_\infty^m(D) \not\subset W_p^m(D)$  for  $p < \infty$  and  $C^\infty(D)$  is not dense in  $W_\infty^m(D)$ . In this case, note that  $C^{\bar{m}}(D) \supset W_\infty^m(D)$ . For  $\bar{m} \geq 1$  and for  $f \in C^{\bar{m}}(D)$ ,

$$(6.7) \quad \begin{aligned} f(x) &= \sum_{|\alpha|=\bar{m}} \frac{(x-y)^\alpha}{\alpha!} f^{(\alpha)}(y) \\ &+ \bar{m} \sum_{|\alpha| \leq \bar{m}} \frac{(x-y)^\alpha}{\alpha!} \int_0^1 s^{\bar{m}-1} [f^{(\alpha)}(x+s(y-x)) - f^{(\alpha)}(y)] \, ds. \end{aligned}$$

This representation is just the first line of the proof of the Sobolev representation given in Section 3, except  $l$  was decreased to  $\bar{m}$  and zero was added in a convenient form. Each term in the second sum can be bounded by

$$\frac{\bar{m}}{\alpha!} |(x-y)^\alpha| \int_0^1 s^{\bar{m}-1} [(1-s)|x-y|]^\theta \, ds |f|_{W_\infty^m(D)}.$$

Multiplying (6.7) by  $\phi(y)$ , integrating with respect to  $y$ , and applying the above bound gives the conclusion for  $p = \infty$  and  $\bar{m} \geq 1$ .

For  $\bar{m} = 0$  replace (6.7) by the trivial relation

$$f(x) = f(y) + [f(x) - f(y)].$$

Then proceed as above.  $\square$

Proposition 6.1, when combined with Theorem 3.1, gives

**THEOREM 6.1.** *Suppose that  $m = \bar{m} + \theta$ , where  $0 < \theta < 1$  and  $\bar{m}$  is a nonnegative integer. Let  $l = \bar{m} + 1$ , and let  $Q^l$  be defined by (3.2). Then there exists a constant  $C = C(n, \phi, d, m)$  such that, for  $1 \leq p \leq \infty$  and  $f \in W_p^m(D)$ ,*

$$(6.8) \quad \|f - Q^l f\|_{W_p^m(D)} \leq C \|f\|_{W_p^m(D)}.$$



**7. Relaxation of Domain Constraints.** The results of Sections 3, 4, 5, and 6 were derived under the assumption that the domain was star-shaped with respect to each point in a ball. In this section we show how this constraint can be weakened. In particular, the previous results can be extended to bounded domains which satisfy the restricted cone condition (see below) that was used in [4], [5]. In addition certain domains which fail to satisfy the restricted cone condition can be treated (for example a slit disk in  $\mathbf{R}^2$ ). The principal result of this section (Theorem 7.1) states roughly that a domain has good approximation properties if it is a finite union of domains with good approximation properties. In [14] Jamet uses a different method to relax geometric constraints associated with polynomial approximation.

First, we remark on the relation between domains which satisfy the restricted cone condition and those which are star-shaped with respect to a ball. A bounded open set  $\Omega$  is said to satisfy the *restricted cone condition* if there exists a finite open cover  $\{O_j\}_{j=1}^J$  of  $\bar{\Omega}$  and a corresponding collection  $\{C_j\}_{j=1}^J$  of truncated right circular cones with vertices at the origin such that if  $x \in \Omega \cap O_j$  then  $x + C_j \subset \Omega$ . The following remark is easily verified.

*Remark 7.1.* If a bounded open set  $\Omega$  satisfies the restricted cone condition then it is the finite union of open sets  $D_j$  each of which is star-shaped with respect to a ball  $B_j$ .

That the converse of this result is not valid is easily seen by considering

$$\Omega = \{re^{i\theta} : 0 < r < 1, 0 < \theta < 2\pi\},$$

where we identify  $\mathbf{C}$  with  $\mathbf{R}^2$ . This domain fails to satisfy the restricted cone condition, but  $\Omega = D_1 \cup D_2$  where

$$D_1 = \{re^{i\theta} : 0 < r < 1, 0 < \theta < 3\pi/2\},$$

$$D_2 = \{re^{i\theta} : 0 < r < 1, \pi/2 < \theta < \pi\}.$$

The domains  $D_j$  are star-shaped with respect to balls  $B_j = \{z : |z - z_j| < 1/4\}$  where  $z_1 = \frac{1}{2}e^{i3\pi/4}$  and  $z_2 = \frac{1}{2}e^{i5\pi/4}$ .  $\square$

For each bounded nonvoid open set  $D$  let  $H(D)$  denote a linear space of functions, and let  $H(D)$  be equipped with two seminorms  $\|\cdot\|_D$  and  $|||\cdot|||_D$ . Suppose that these spaces and seminorms have the following properties:

(a) The restriction of each element of  $H(D_1 \cup D_2)$  to  $D_1$  is in  $H(D_1)$ .

(b) For each  $f \in H(D_1 \cup D_2)$ ,

$$\|f\|_{D_1 \cup D_2} \leq \|f\|_{D_1} + \|f\|_{D_2} \leq 2\|f\|_{D_1 \cup D_2}$$

$$\text{and } |||f|||_{D_1} + |||f|||_{D_2} \leq 2|||f|||_{D_1 \cup D_2}.$$

(c)  $\mathcal{P}_\infty \subset H(D)$ .

(d) If  $P \in \mathcal{P}_\infty$  and  $\|P\|_{D_1} = 0$ , then  $\|P\|_{D_1 \cup D_2} = 0$ .

In the use of the results of this section,  $\|f\|_D$  will be a finite sum of terms of the form  $\|f^{(\alpha)}\|_{L_p(D)}$  and  $|||f|||$  will include in addition terms of the form

$$\|P(\partial/\partial x)f\|_{L_p(D)} \text{ and } |f^{(\gamma)}|_{w_p^\theta(D)}.$$

**THEOREM 7.1.** *Suppose that  $\Omega = \bigcup_{j=1}^N D_j$  is connected and that each  $D_j$  is a bounded nonvoid open set. Let  $\mathcal{P}$  be a finite-dimensional subspace of  $\mathcal{P}_\infty$ , and suppose that there exist  $Q_j$  and  $C_j$  for  $j = 1, \dots, N$  such that, for  $f \in H(D_j)$ ,  $Q_j f \in \mathcal{P}$  and*

$$(7.1) \quad \|f - Q_j f\|_{D_j} \leq C_j \|f\|_{D_j}.$$

*Then there exists  $C_0$  such that, for  $j = 1, \dots, N$  and  $f \in H(\Omega)$ ,*

$$(7.2) \quad \|f - Q_j f\|_\Omega \leq C_0 \|f\|_\Omega.$$

*Proof.* It suffices to consider the case  $N = 2$ , since the general case follows easily by induction. Let  $B = D_1 \cap D_2$ ;  $B \neq \emptyset$  since  $\Omega$  is connected. By properties (b) and (d) above, the seminorms  $\|P\|_{D_1} + \|P\|_{D_2}$  and  $\|P\|_B$  on  $P \in \mathcal{P}$  have the same kernel. Using the equivalence of norms on the corresponding quotient space yields

$$(7.3) \quad \|P\|_{D_1} + \|P\|_{D_2} \leq C \|P\|_B \quad \text{for all } P \in \mathcal{P}$$

for some constant  $C = C(D_1, D_2, \mathcal{P})$ .

Suppose that  $f \in H(\Omega)$  and that  $P_j = Q_j f$  for  $j = 1, 2$ . Note that

$$\|f - P_1\|_\Omega \leq \|f - P_1\|_{D_1} + \|f - P_2\|_{D_2} + \|P_2 - P_1\|_{D_2},$$

using property (b) and the triangle inequality. By (7.3),

$$\|P_2 - P_1\|_{D_2} \leq C \|P_2 - P_1\|_B \leq C [\|P_2 - f\|_B + \|f - P_1\|_B].$$

Combining with the previous inequality, applying (7.1), and using property (b) yields

$$\begin{aligned} \|f - P_1\|_\Omega &\leq (1 + C)(C_1 \|f\|_{D_1} + C_2 \|f\|_{D_2}) \\ &\leq (1 + C) \max\{C_1, C_2\} \|f\|_\Omega. \quad \square \end{aligned}$$

**Remark 7.2.** In those cases in which the norm  $\|\cdot\|_D$  is translation invariant, as is the case for all the Sobolev-type seminorms used so far in this paper, the constant  $C$  in (7.3) can be taken to depend only on  $\underline{d}$  and  $\bar{d}$  instead of  $D_1$  and  $D_2$ , where  $\underline{d} > 0$  is such that some ball of radius  $\underline{d}$  is contained in  $D_1 \cap D_2$  and  $\bar{d} = \text{diam}(D_1 \cup D_2)$ .  $\square$

**Remark 7.3.** It follows from Theorem 7.1 that Theorems 3.2, 4.2, 5.1, and 6.1 hold if  $D$  is any connected open set that is the union of a finite collection of domains that are star-shaped with respect to balls. In Theorems 3.2, 4.2, and 6.1, one chooses  $Q$  to be defined with respect to any ball contained in  $D$ . Note that Theorem 5.1 still holds because Theorem 7.1 does not require the mapping  $Q_j$  to be linear or even continuous; hence, we may define  $Q_j$  by taking anything reasonably close to the infimum in Theorem 5.1.  $\square$

**8. Examples.** This section contains four simple examples that are based on the results of Sections 4, 5, and 6. The purpose here is to show how the refinements in those sections yield results that would not be easily derived by results based on complete polynomial approximation or on more restrictive tensor product results. First the results of Section 4 are used to show an error bound for approximation by polynomials that are constant in one variable and linear in another. Next, the results of

Section 5 are used to show how well harmonic polynomials can approximate harmonic functions. In the third example, the results of Section 6 are used to bound the interpolation error in a case in which the function being interpolated does not have enough derivatives to be able to apply the theorems of Section 3. The fourth example shows how triangles with curved edges can be treated using our results.

*Example 1.* In this example we consider approximation in two variables by polynomials that are constant in one variable and linear in the other. One interesting question in this context is whether differentiation of the approximating polynomial in the direction in which it is linear gives a good constant approximation to the derivative of the function being approximated. The commutativity of the operator  $Q^A$  with the differentiation operator allows an affirmative conclusion.

Let  $h$  be a positive parameter, and let  $D_h = (0, h) \times (0, h)$ . Take  $\beta = (1, 0)$  and consider, for  $f \in C^\infty(\mathbb{R}^2)$ ,

$$(8.1) \quad \eta(f; h) = \inf \left\{ \|f - P\|_{L_2(D_h)} + \left\| \left( \frac{\partial}{\partial x} \right)^\beta (f - P) \right\|_{L_2(D_h)} : P \in \mathcal{P} \right\},$$

where  $\mathcal{P} = \{a + bx_1 : a, b \in \mathbb{R}\}$ .

For any function  $g$  defined on  $D_h$  let  $\tilde{g}(x_1, x_2) = g(hx_1, hx_2)$  be defined on  $D_1$ . Note that

$$(8.2) \quad \left\| \left( \frac{\partial}{\partial x} \right)^\alpha (f - P) \right\|_{L_2(D_h)} = h^{1-|\alpha|} \left\| \left( \frac{\partial}{\partial x} \right)^\alpha (\tilde{f} - \tilde{P}) \right\|_{L_2(D_1)}.$$

Thus, since  $P \in \mathcal{P}$  if and only if  $\tilde{P} \in \mathcal{P}$ ,

$$\eta(f, h) = \inf \left\{ h \|\tilde{f} - \tilde{P}\|_{L_2(D_1)} + \left\| \left( \frac{\partial}{\partial x} \right)^\beta (\tilde{f} - \tilde{P}) \right\|_{L_2(D_1)} : \tilde{P} \in \mathcal{P} \right\}.$$

Fix  $\phi \in C_0^\infty(B)$  such that  $\int \phi = 1$ , where  $B$  is a ball contained in  $D_1$ . Take  $A = \{(0, 1), (2, 0)\}$  and let  $Q = Q^A$  be defined as in Section 4. Then, by Theorem 4.2,

$$(8.3) \quad \begin{aligned} h \|\tilde{f} - Q\tilde{f}\|_{L_2(D_1)} &\leq hC [\|\tilde{f}^{(0,1)}\|_{L_2(D_1)} + \|\tilde{f}^{(2,0)}\|_{L_2(D_1)}], \\ \left\| \left( \frac{\partial}{\partial x} \right)^\beta (\tilde{f} - Q\tilde{f}) \right\|_{L_2(D_1)} &\leq C [\|\tilde{f}^{(1,1)}\|_{L_2(D_1)} + \|\tilde{f}^{(2,0)}\|_{L_2(D_1)}]. \end{aligned}$$

Adding the two inequalities in (8.3) and applying (8.2) with  $P \equiv 0$ , we see that for  $0 < h \leq 1$

$$(8.4) \quad \eta(f; h) \leq Ch [\|f^{(0,1)}\|_{L_2(D_h)} + \|f^{(1,1)}\|_{L_2(D_h)} + \|f^{(2,0)}\|_{L_2(D_h)}].$$

It is interesting to note that if we had restricted ourselves to the *direct* application of the results of Section 3 we would have not been able to show that  $\eta(f; h) = O(h)$  since the largest class of complete polynomials contained in  $\mathcal{P}$  is  $\mathcal{P}_1$  and  $\eta(f; h) \geq \|(\partial/\partial x)f\|_{L_2(D)}$  if  $\mathcal{P}$  is replaced by  $\mathcal{P}_1$  in the inf. Our results are related to those used by Ewing [10] in deriving a similar cross derivative approximation bound.

To further illustrate the possible uses of Theorem 4.2, suppose that  $q, p_1, p_2 \in [1, \infty]$ ,  $f^{(0,1)} \in L_{p_1}(D_h)$ , and  $f^{(2,0)} \in L_{p_2}(D_h)$ , where  $1/q - 1/p_1 + 1/2 > 0$ . Then

$$\inf \{ \|f - P\|_{L_q(D_h)} : P \in \mathcal{P} \} \leq Ch^{1-2/p_1+2/q} \|f^{(0,1)}\|_{L_{p_1}(D_h)} + h^{2-2/p_2+2/q} \|f^{(2,0)}\|_{L_{p_2}(D_h)}.$$

In general, direct application of results of Section 3 would not yield such a bound.  $\square$

*Example 2.* To construct an example of the use of Theorem 5.1, let  $P_{j+1}(x) = x^{(r+1-j,i)}$ , for  $j = 0, \dots, r+1$ , and let  $P_{r+3}(x_1, x_2) = x_1^2 + x_2^2$ . The set  $K$  consists of all harmonic polynomials of degree less than or equal to  $r$ . If we proceed exactly as in the previous example, we see that for all  $f \in C^\infty(\mathbb{R}^2)$

$$(8.5) \quad \inf_{P \in K} \|f - P\|_{L_2(D_h)} \leq C \left[ h^{r+1} \sum_{|\alpha|=r+1} \|f^{(\alpha)}\|_{L_2(D_h)} + h^2 \|\Delta f\|_{L_2(D_h)} \right].$$

We can conclude that if  $f$  is harmonic on  $D_h$  then it can be approximated by elements of  $K$  with an error  $Ch^{r+1}$ .  $\square$

*Example 3.* To give an application of Theorem 6.1 we consider the question of bounding the error in an interpolation process. Suppose that  $\Omega$  is a bounded domain in  $\mathbb{R}^2$  with a polygonal boundary and that  $\mathcal{F}$  is a family of triangulations of  $\Omega$ . For  $T \in \mathcal{F}$ , let  $h = h(T) = \max_{T \in \mathcal{T}} \text{diam}(T)$ . Denote by  $M_T = M_0(1, T)$  the space of functions that are continuous on  $\bar{\Omega}$  and linear (affine) on each  $T \in \mathcal{T}$ . Assume that there is a  $\rho < \infty$  such that for all  $T \in \mathcal{F}$ ,  $T \in \mathcal{T} \Rightarrow (\text{diam}(T))^2 / \text{area}(T) \leq \rho$ ; this says that the triangles do not degenerate.

For any function  $f \in C(\bar{\Omega})$  let  $If = I_T f$  be the element of  $M_T$  which agrees with  $f$  at each vertex of  $T$ ; i.e.,  $If$  is the piecewise linear interpolant of  $f$ . It is well known (and follows easily from Theorem 3.2) that if  $f \in W_2^2(\Omega)$ , then

$$\|f - If\|_{L_2(\Omega)} \leq Ch^2 \|f\|_{W_2^2(\Omega)}.$$

However, if  $0 < \epsilon < 1$  and if  $f \in W_2^{1+\epsilon}(\Omega)$ , the results of Section 3 give no error bound. A natural approach would be to try to use the theory of interpolation of Banach spaces and use results for  $W_2^2$  and  $W_2^1$ ; however, this fails because  $If$  is not defined on  $W_2^1$  since the elements of this space are not in general continuous.

Define  $T_R = \{(x_1, x_2) : x_1 > 0, x_2 > 0, x_1 + x_2 < 1\}$ . For  $T$  a triangle in  $\mathcal{T} \in \mathcal{F}$  let  $A$  be an affine map taking  $T_R$  onto  $T$ . Assume, without loss of generality, that  $A$  is linear, and note that  $\|A\| \leq Ch$ . For a function  $\tilde{g} \in C(\bar{T}_R)$  define  $\tilde{I}\tilde{g}$  to be the affine function of  $x$  that agrees with  $\tilde{g}$  at the vertices of  $T_R$ ; i.e.,

$$\tilde{I}\tilde{g}(x_1, x_2) = \tilde{g}(0, 0)(1 - x_1 - x_2) + \tilde{g}(1, 0)x_1 + \tilde{g}(0, 1)x_2.$$

Note that  $(\tilde{f} - \tilde{I}\tilde{f})(x) = (f - If)(Ax)$ , where  $\tilde{f}(x) \equiv f(Ax)$ . Thus

$$(8.6) \quad \|f - If\|_{L_2(T)}^2 = |\det A| \|\tilde{f} - \tilde{I}\tilde{f}\|_{L_2(T_R)}^2.$$

Next note that for any  $P \in \mathcal{P}_2$

$$\|\tilde{f} - \tilde{I}\tilde{f}\|_{L_2(T_R)} = \|\tilde{f} - P - \tilde{I}(\tilde{f} - P)\|_{L_2(T_R)},$$

since  $\tilde{I}$  is a linear map which reproduces polynomials in  $\mathcal{P}_2$ . Letting  $\|\tilde{I}\|$  denote the norm of  $\tilde{I}$  as a map of  $W_2^{1+\epsilon}(T_R)$  into  $L_2(T_R)$  (which is finite by Sobolev's inequality) we see that

$$\|\tilde{f} - \tilde{I}\tilde{f}\|_{L_2(T_R)} \leq (1 + \|\tilde{I}\|) \inf_{P \in \mathcal{P}_2} \|\tilde{f} - P\|_{W_2^{1+\epsilon}(T_R)}.$$

From Theorem 6.1 and the fact that  $|P|_{W_2^{1+\epsilon}} = 0$  for  $P \in \mathcal{P}_2$  it follows that

$$(8.7) \quad \|\tilde{f} - \tilde{I}\tilde{f}\|_{L_2(T_R)} \leq C|\tilde{f}|_{W_2^{1+\epsilon}(T_R)}.$$

To estimate the right-hand side of (8.7) represent  $A$  as a  $2 \times 2$  matrix that acts on column vectors  $(x_1, x_2)^T$  and then note that

$$\nabla \tilde{f} = \begin{pmatrix} \tilde{f}^{(1,0)} \\ \tilde{f}^{(0,1)} \end{pmatrix} = A^* \begin{pmatrix} f^{(1,0)} \\ f^{(0,1)} \end{pmatrix} = A^* \nabla f,$$

where  $\tilde{f}^{(\alpha)}$  and  $f^{(\alpha)}$  are evaluated at  $x$  and  $Ax$ , respectively. Thus

$$\begin{aligned} & \iint_{T_R \times T_R} \frac{|\nabla \tilde{f}(x) - \nabla \tilde{f}(y)|^2}{|x - y|^{2+2\epsilon}} dx dy \\ &= |\det A|^{-2} \iint_{T \times T} \frac{|A^*(\nabla f(x) - \nabla f(y))|^2}{|x - y|^{2+2\epsilon}} \left( \frac{|x - y|}{|A^{-1}(x - y)|} \right)^{2+2\epsilon} dx dy \\ &\leq \frac{\|A\|^{4+2\epsilon}}{|\det A|^2} \iint_{T \times T} \frac{|\nabla f(x) - \nabla f(y)|^2}{|x - y|^{2+2\epsilon}} dx dy. \end{aligned}$$

From this result, (8.6) and (8.7) we see that

$$\|f - If\|_{L_2(T)}^2 \leq C \left[ \frac{\|A\|^2}{|\det A|} \right] h^{2+2\epsilon} \iint_{T \times T} \frac{|\nabla f(x) - \nabla f(y)|^2}{|x - y|^{2+2\epsilon}} dx dy.$$

Sum this result over triangles and use the nondegeneracy of  $F$  to bound the term in brackets to obtain

$$\begin{aligned} (8.8) \quad \|f - If\|_{L_2(\Omega)} &\leq Ch^{1+\epsilon} \left( \sum_{T \in \mathcal{T}} \iint_{T \times T} \frac{|\nabla f(x) - \nabla f(y)|^2}{|x - y|^{2+2\epsilon}} dx dy \right)^{1/2} \\ &\leq Ch^{1+\epsilon} |f|_{W_2^{1+\epsilon}(\Omega)}. \end{aligned}$$

If we had not needed to estimate the interpolation error but merely the error in the best possible approximation in  $M_T$  a bound like (8.8) could be obtained by interpolating between  $L_2(\Omega)$  and  $W_2^2(\Omega)$ . However, it is frequently the case that one needs

to know how well a function that vanishes on the boundary can be approximated by function spaces that vanish on the boundary. In such cases bounds like (8.8) extend the error estimates to their natural lower limits. One such example can be found in Douglas, Dupont, Percell and Scott [8].  $\square$

*Example 4.* We now show how the above results can be applied to certain families of curved domains. Suppose  $\Omega$  is a bounded domain in  $\mathbf{R}^2$  with smooth boundary  $\partial\Omega$ . Let  $\mathcal{F}$  be a family of triangulations of  $\Omega$  having straight interior edges and (possibly) curved edges lying on  $\partial\Omega$ , and suppose that  $\mathcal{F}$  satisfies the nondegeneracy assumption of Example 3:

$$\sup_{T \in \mathcal{F}} \sup_{T \in \mathcal{T}} (\text{diam}(T))^2 / \text{area}(T) \leq p < \infty.$$

Such families of triangulations were considered in [17], where approximation properties for the boundary triangles having a curved side were derived in a very complicated way. The main difficulty is that now there is no fixed reference triangle, but rather a family of reference domains. For each  $T \in \mathcal{T} \in \mathcal{F}$ , define an affine mapping by sending the vertices of  $T$  onto the set  $\{(0, 0), (1, 0), (0, 1)\}$ , and let the image of  $T$  be denoted  $T_R$ . Again define  $h(T) = \max_{T \in \mathcal{T}} \text{diam}(T)$  for all  $T \in \mathcal{F}$ . For  $h_0$  sufficiently small (depending only on  $\Omega$  and  $\rho$ ), if  $h(T) \leq h_0$  and  $T \in \mathcal{T}$ , then  $T_R$  is contained in the disc  $\{|x| \leq 2\}$  and is star-shaped with respect to the disc  $\{|x - x_0| \leq 1/8\}$ , where  $x_0 = (1/4, 1/4)$ . Thus, the above approximation results apply to  $T_R$  and, via the affine mapping, to each  $T$ , with the constant  $C$  in the estimates depending only on  $\Omega$  and  $\rho$  (as well as the degree and type of polynomial approximation).  $\square$

Department of Mathematics  
University of Chicago  
Chicago, Illinois 60637

Applied Mathematics Department  
Brookhaven National Laboratory  
Upton, New York 11973

1. S. AGMON, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, Princeton, N. J., 1965.
2. R. ARCANGELI & J. L. GOUT, "Sur l'évaluation de l'erreur d'interpolation de Lagrange dans un ouvert de  $\mathbf{R}^n$ ," *RAIRO Analyse Numérique*, v. 10, 1976, pp. 5–27.
3. A. BERGER, R. SCOTT & G. STRANG, "Approximate boundary conditions in the finite element method," *Symposia Mathematica* X, Academic Press, New York, 1972, pp. 295–313.
4. J. H. BRAMBLE & S. R. HILBERT, "Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation," *SIAM J. Numer. Anal.*, v. 7, 1970, pp. 112–124.
5. J. H. BRAMBLE & S. R. HILBERT, "Bounds for a class of linear functionals with applications to Hermite interpolation," *Numer. Math.*, v. 16, 1971, pp. 362–369.
6. V. I. BURENKOV, "Sobolev's integral representation and Taylor's formula," *Trudy Mat. Inst. Steklov.*, v. 131, 1974, pp. 33–38.
7. P. G. CIARLET & C. WAGSCHAL, "Multipoint Taylor formulas and applications to the finite element method," *Numer. Math.*, v. 17, 1971, pp. 84–100.
8. J. DOUGLAS, JR., T. DUPONT, P. PERCELL & R. SCOTT, "A family of  $C^1$  finite elements with optimal approximation properties for various Galerkin methods for 2nd and 4th order problems," *RAIRO Analyse Numérique*, v. 13, 1979, pp. 227–255.
9. T. DUPONT & R. SCOTT, "Constructive polynomial approximation in Sobolev spaces," *Recent Advances in Numerical Analysis* (C. de Boor and G. Golub, Eds.), Academic Press, New York, 1978.

10. R. E. EWING, "Alternating direction Galerkin methods for some third and fourth order equations." (To appear.)
11. A. FRIEDMAN, *Partial Differential Equations*, Holt, Rinehart, and Winston, New York, 1969.
12. P. GRISVARD, "Behavior of solutions of an elliptic boundary value problem in polygonal or polyhedral domains," *Numerical Solution of Partial Differential Equations—III (Synspade 1975)* (B. Hubbard, Ed.), Academic Press, New York, 1976, pp. 207–274.
13. L. HÖRMANDER, *Linear Partial Differential Operators*, Springer-Verlag, Berlin, 1963.
14. P. JAMET, "Estimation de l'erreur d'interpolation dans un domaine variable et application aux éléments finis quadrilatéraux dégénérés," *Méthodes Numérique en Mathématiques Appliquées*, Presses de l'Université de Montréal, 1977, pp. 55–100.
15. J. MEINGUET, "Structure et estimations de coefficients d'erreurs," *RAIRO Analyse Numérique*, v. 11, 1977, pp. 355–368.
16. C. B. MORREY, *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, Berlin, 1966.
17. R. SCOTT, "Interpolated boundary conditions in the finite element method," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 404–427.
18. S. L. SOBOLEV, *Applications of Functional Analysis in Mathematical Physics*, Transl. Math. Monographs, Vol. 7, Amer. Math. Soc., Providence, R. I., 1963.
19. E. M. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, N. J., 1970.
20. B. L. VAN DER WAERDEN, *Modern Algebra II*, 2nd ed., Ungar, New York, 1950.
21. A. ZYGMUND, *Trigonometric Series* (2 volumes), 2nd ed., Cambridge Univ. Press, London, 1959.